Hollywood Workers vs Tech: In Theory and In the News

by

Christian Cmehil-Warn

B.S. Economics, University of Missouri, 2020
B.S. Statistics, University of Missouri, 2020

Submitted to the Institute for Data, Systems, and Society and the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

MASTER OF SCIENCE IN TECHNOLOGY AND POLICY

and

MASTER OF SCIENCE IN COMPUTER SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

© 2024 Christian Cmehil-Warn. This work is licensed under a CC BY-NC-ND 4.0 license.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Christian Cmehil-Warn
Institute for Data, Systems, and Society
Department of Electrical Engineering and Computer Science
May 10, 2024

Certified by: Jason Jackson
Associate Professor of Political Economy and Urban Planning
Thesis Supervisor

Certified by: Manish Raghavan
Assistant Professor of Information Technology
Thesis Supervisor

Accepted by: Frank R. Field III
Senior Research Engineer, Sociotechnical Systems Research Center
Interim Director, Technology and Policy Program

Accepted by: Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Hollywood Workers vs Tech: In Theory and In the News

by

Christian Cmehil-Warn

Submitted to the Institute for Data, Systems, and Society and the
Department of Electrical Engineering and Computer Science
on May 10, 2024 in partial fulfillment of the requirements for the degrees of

MASTER OF SCIENCE IN TECHNOLOGY AND POLICY

and

MASTER OF SCIENCE IN COMPUTER SCIENCE

ABSTRACT

The 2023 SAG-AFTRA and WGA strikes on Hollywood were notable because of their explicit ties to technology and labor's changing relationships. In particular, disputes around using generative AI in the workplace were widely reported in the news. This thesis examines the Hollywood strikes in two parts. The first part takes a political economy approach to examine the underlying causes of these changes in technology-labor relations. In particular, the thesis argues that an industry shift to distribution via streaming services alongside increased vertical integration brought about new imperatives to production and exponentially increased levels of data capture, enabling the labor conditions that led to the strike. Theories of creative labor and technology-labor relations are used to describe the tensions. The resulting SAG-AFTRA and WGA collective bargaining agreements are then examined within these framings.

The second part of the thesis quantitatively explores the relationship between news media (which its own complex relationship with technology) and the Hollywood strikes using natural language processing techniques. Sentiment analysis and sentence embeddings are used to quantify and compare news articles across different characteristics. The results of the analysis are inconclusive.

Thesis supervisor: Jason Jackson
Title: Associate Professor of Political Economy and Urban Planning

Thesis supervisor: Manish Raghavan
Title: Assistant Professor of Information Technology

Acknowledgments

My term here at MIT has coincided with a period of great change and challenge for the institute. So first, I want to thank my fellow organizers, bargaining committee members, and workers in the MIT Graduate Student Union. The opportunities I had with MITGSU have given meaning to my time here at MIT and changed the course of my life. I've met so many wonderful people through the union who have challenged me and helped me grow as an organizer, thinker, and person. You all have had my back when navigating the bureaucratic challenges of an institution like MIT. I also want to thank the United Electrical, Radio, and Machine Workers of America, our national affiliate, for giving us the support necessary to build our union and the trust to truly make it our own as workers.

This thesis would not have happened without my thesis advisor, Professor Jason Jackson, and thesis reader, Professor Manish Raghavan, for taking on the challenge of helping refine my smattering of ideas into a comprehensible paper. I'd also like to thank all of the professors who funded me, allowing me to continue in this program. Rui-Jie and Deepika have given me so much-needed reassurance to keep going in this program that I can't imagine finishing without them. Also, thank you to the National Telecommunications and Information Administration; Office for Policy Analysis and Development for giving me a phenomenal summer internship and to Travis and Ona for helping me narrow down my initial thesis ideas. Thanks to Nishad for the section title "Lights, Camera, Collective Action".

I would not have reached this point without the tremendous support of faculty and staff from the University of Missouri, my alma mater. People like Rachel Harper, Rachel Newman, Kristina Bradley, Megan Boyer, Tim Parshall, Thomas Kane, Jay Sexton, and countless others helped me draft applications, find opportunities, and apply for fellowships, along with their emotional support. I'd also like to thank my high school science teacher, Mrs. Karlish, for always being in my corner.

Reading and listening to rigorous and thorough criticism of technologies has been essential to my intellectual development. The podcast *This Machine Kills* has been a boundless well of inspiration for me, including for this thesis, which was inspired by Episode 292, where I first saw Delfanti [1] put in the context of the Hollywood strikes. The hosts, Ed, Jathan, and Jeremy, are fantastic, and it's been great being able to talk with them.

Last but certainly not least, I am in endless debt to my friends and family for always being a rock in an absolutely wild time of my life. Bekah, Jonah, Jesse, Amelia, Joe, and Jack deserve honorary degrees for how much they listen to my rambling. My moms, Paula and Sueanne, are two of the best parents anyone could ask for. They are kind, loving, supportive, and always happy to take a call when I'm panicking or stressed or lonely or just want to chat. I love you two with all my heart.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Writers and actors are gig workers in the most literal sense of the word. Instead of continuous employment, they work discrete, finite jobs (“gigs“) on specific projects. Recently, gig work (and its derivatives “gigification“ and “gig economy“) has become a pejorative description of the labor models and worker misclassification innovated by tech companies. Many formerly stable professions have been subsumed into the “gig economy,“ resulting in worker precarity.[1] With changing industry structures and technology, actors and writers were concerned about facing similar conditions; however, writers and actors can collectively push back against such changes through their labor unions, the Writer’s Guild of America (WGA) and the Screen Actors Guild-American Federation of Television and Radio Artists (SAG-AFTRA) [2], [3].

SAG and WGA are different from most unions in America because they are craft unions, also called guilds. They represent workers of a specific trade to various employers, unlike the more common industrial union model, which represents workers at one particular firm. In other words, instead of representing actors at a specific studio, they represent actors across studios by setting base terms that studios agree to either individually or collectively, as the members of the Alliance of Motion Picture and Television Producers (AMPTP) do. SAG and WGA work to create stable benefits (such as health insurance) and payment for workers who frequently switch between jobs and employers [2].

It’s important to note that SAG and WGA have significantly different member compositions. SAG-AFTRA has a heterogeneous membership of about 160,000 members that include lead actors, background actors, dancers, stunt performers, singers, puppeteers, and more. Contrastingly, WGA has about 11,000 members mainly composed of writers. Both unions also have different governance structures, the details of which are out of the scope of this paper, but both unions resolve issues like authorizing a strike or a tentative agreement with a membership vote. WGA has historically been more willing to take strike action, with their last strike in 2007, compared to SAG, who last went on strike in 1980 [4], [5].

[1] In this paper, gig work and its derivatives will refer to the modern definition unless otherwise specified.

1.1 Lights, Camera, Collective Action: The 2023 Hollywood Strikes

Due to changing conditions and an obstinate AMPTP, SAG and WGA went on strike in 2023. The last time that both unions simultaneously struck their labor was 63 years prior. The WGA went on strike first, with the membership authorizing the strike on April 17 with 97.9% voting in favor, the highest margin in WGA's history, and then starting the strike on May 2. SAG responded by authorizing a strike on June 5, two days before they *started* bargaining with the AMPTP and went on strike on July 13.[6]

Both strikes were resolved in the same order, with WGA reaching a tentative agreement on September 25 (146 days on strike) and SAG reaching a tentative agreement on November 8 (118 days on strike). The strikes had a significant effect on Hollywood by halting production and delaying releases for a wide range of movies and TV shows.[6] The next section of the paper analyses why these strikes happened and how the tentative agreements left issues around technology resolved or unresolved, followed by a quantitative analysis of news reporting about the strikes.

Chapter 2

Striking the Set: The Hollywood Guilds' Collective Action Against Technology

2.1 Establishing Shot: Conditions Going Into the 2023 Hollywood Strikes

2.1.1 Vertical Integration and the Imperative for Content

Due to its capital-intensive nature, movie production started in the 1920s as a hyper-consolidated industry with eight studios controlling 95% of films, from production to distribution to exhibition. However, this vertical integration was disrupted in the 1940s by antitrust lawsuits, television proliferation, and an overall decrease in consumers' non-essential spending, resulting in a 74% decrease in profits from 1946 to 1956 among the ten largest studios. Studios were acquired by non-media corporations for their content libraries, as opposed to their production capacity. In contrast, television, whose profits quickly overtook film, began as a much more regulated industry with limitations on broadcast station ownership and network-produced programming revenues. In the 1980s, cable distribution enabled an exponential increase in channels, which was used to justify widespread deregulation in television. This deregulation resulted in television networks and film studios consolidating into large media conglomerates. [7], [8]

While this structural shift was solidifying in the 2000s, services like Netflix, Hulu, and Amazon Prime caused a film distribution revolution with a new direct-to-consumer model, streaming. Instead of going through television networks, movie theaters, or physical media, consumers could subscribe to these streaming services to see content on demand. Originally, these services did not produce their own content but instead aggregated content from traditional media distributors into a more modern interface with features like personalized recommendations and on-demand programming (as opposed to strictly linear scheduling). Consumers started "cutting the cable" and moving to these new services, with 59% of Americans dropping cable services by 2019. [7], [8]

Television advertising was a significant revenue source for media companies. However, advertisers shifted to the Internet simultaneously as consumers left for streaming services. Traditional media conglomerates had been largely outmaneuvered in their vertical integra-

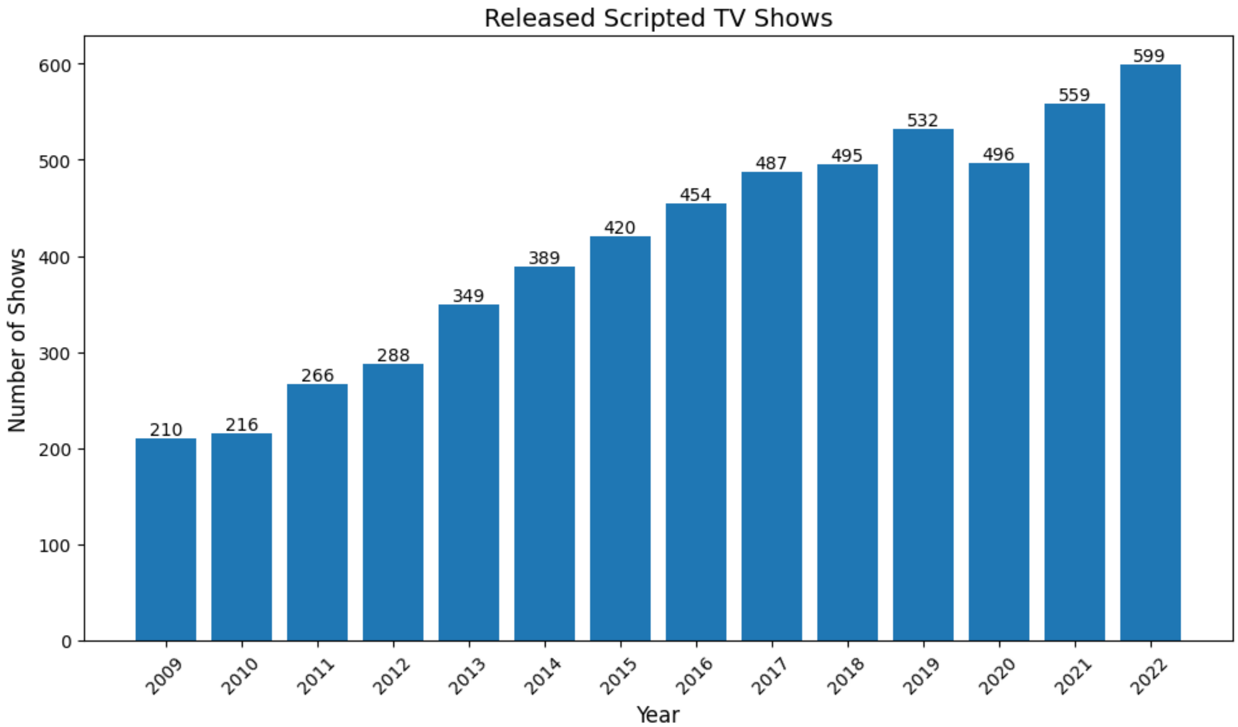


Figure 2.1: Data from Porter [10]

tion, with streaming services taking over the "last mile" of media distribution. In response, traditional media companies shifted from a revenue base of advertising to a base of consumer spending and stopped licensing popular, existing content with 3rd party streaming services. Instead, they started their own walled gardens that attempted to entice customers with exclusive content, both old and new. This change in strategy has resulted in pre-streaming-era, popular media being licensed in multi-million-dollar deals, as well as a major push for new content, as seen in Figure 2.1. [8], [9]

Movie production has also shifted due to the rise of streaming. The medium-budget movie ($20mil-$100mil) has gone by the wayside, decreasing the overall quantity of movies produced as theater revenues fell, even before the COVID-19 pandemic. Instead, high-budget movies ballooned in cost, making them riskier bets. Low-budget films held constant. [11]

Streaming also enables conglomerates to have this increased spending on movies and television shows be informed by unprecedented amounts of data generated from streaming services. Netflix was an early practitioner of this new form of data-driven decision-making, ordering and tweaking original television content based on the minute details of customers' viewing behaviors. This practice has led to more direct involvement of studio executives in production decisions that would normally be left to production workers. [8]

Streaming technology and vertical integration have created a new imperative for content, requiring media production to significantly increase its outputs to make this new state profitable for traditional media conglomerates. As one Marvel Studios (owned by Disney) executive notes about the early days of Disney+: "I mean, frankly, in all honesty, there was a mandate to kind of create as much as we could for Disney+ as quickly as we could."

[12] Streaming has significantly changed the relationships writers and actors have with the outputs of their labor and incentivized conglomerates to find new ways to increase output and to drive detailed production decisions using streaming's troves of data in the pursuit of decreasing risk of losses.

2.1.2 Stream-a-nomics: Streaming and Residuals Going Into the Strike

The transition from cable and theatrical media release to streaming services, a significant issue in the 2007-2008 writers strike [13], upended the compensation structures of WGA and SAG. Alongside upfront payment, writers and non-background actors would receive residuals, additional payment after production calculated based on the continued use of the media. In theory, Hollywood workers could rely on this income to smooth over financially rough periods between jobs. [14], [15]

WGA residuals are based initially on the credit a writer receives for a script. For example, a writer with "Written by" credit nets 100% of the residual calculation while a "Story by" credited writer receives 25% of the calculated residual. From there, there are two types of residual calculations: Revenue-based residuals are calculated based on a distributor's gross revenue, which includes theatrical-release movies (only post-theatrical release revenue) and non-"high-budget" made-for-streaming TV [1] among other types of media. Fixed residuals cover "high-budget" made-for-streaming TV, Network TV, and other forms of syndicated television. A set base residual amount is used for calculations that consider things like the number of times a show airs (for traditional TV) and use in foreign markets, among other details. Specifically for "high-budget" made-for-streaming TV, the base residual is modified based on how many subscribers a service has and is decreased the longer the program is on the service. There's also a separate payment for streaming in foreign markets that (pre-2023) does not scale with the amount of subscribers. [16]–[19]

For SAG members, residuals work similarly, except for body doubles and background actors, who do not receive residuals. The main differences lie in the variety of payment structures (given the wide range of SAG-covered work) and the minutiae of fixed residuals calculations. For example, instead of a set base residual amount, residuals are based on a SAG member's initial compensation and have a ceiling. The residual structure for "high-budget" made-for-streaming TV has the same modifiers as the WGA residuals. SAG also has a separate payment for streaming on foreign services that (pre-2023) does not scale with the amount of subscribers. [20]

Notably, streaming represents a capture of what was once somewhat observable data. While exact viewership numbers pre-streaming were challenging to collect, simply by looking at the programming schedules for theaters and TV channels, workers could use this data as leverage to demand payment from studios. As former SAG president Barry Gordon puts it: “When it was three networks [on television], it was pretty easy to know where that show was playing anytime it was playing.“ [15] Streaming, however, is entirely centralized. Streaming has no regularly scheduled programming to reference, so the number of times a piece of

[1] Pre-2023, a 20-35 minute show with a budget above $1.3 million or a 36-65 minute program with a budget above $1.7 million would be considered "high-budget"

media is run is nearly centrally controlled and lets studies have more leverage in bargaining. [21] The pre-2023 residual structures around streaming content reflect this as there's no mechanism for additional payment based on the success of a show. For example, if a movie or television show was successful in the pre-streaming era, it would be redistributed with some frequency, whether through re-runs or DVD sales, and residual payments would reflect that. However, if a streaming show is the most watched show on the platform for weeks, workers would see no additional compensation.

For television writers in particular, streaming's imperative for increased content production has changed the working environment of writers. Previously, writer's rooms would be groups of writers hired on for a season's entire production period. The writer's room allowed beginning writers to be exposed to the whole production process of a show, creating opportunities for education and advancement. Notable writers, such as George R.R. Martin, observe how being a part of a writer's room meant being a part of the entire production process where they learned the necessary skills to advance in their careers. [22] In contrast, in the streaming era, studios developed a practice writer's derisively call "mini-rooms". Mini-rooms typically consist of fewer writers who write up a whole television season and then are let go before production begins, with only the showrunner staying on for production if the season is picked up. These new employment structures have direct consequences on a writer's financial stability as well as opportunities for advancement.

2.1.3 ChatGP-TV: Generative AI Enters the Scene

Generative AI is dependant on training materials created by workers, whether screenplays written by writers [23] or detailed data about the likeness of actors [24]. Writers were pre-emptively concerned about generative AI being used to reduce their autonomy and the credit they receive on projects. In bargaining with WGA and SAG-AFTRA, studios demonstrated their willingness to engage in these practices by initially rejecting a WGA proposal to prevent AI-generated work from counting as "source material" for a script, which would affect the credits a writer would receive. [23]. Notably, as mentioned earlier, credit on a script is not just a point of pride but a mechanism for how residuals are distributed. Generative AI models receiving credit would create a scenario where a tool created from writers' outputs but leased by a studio would cut into a writer's earnings on a script.

Simultaneously, studios rejected proposals that would prevent writers from being forced to use generative AI. [23] Workers describe how generative AI models could potentially create more work as they could be required to improve poor-quality, AI-generated scripts or that a studio exec could "rewrite " a scene using generative AI. Writers were largely not concerned with replacement but with generative AI imposing conditions that make their jobs worse by attacking their pay and autonomy. [25]

In contrast, actors, particularly background actors, were actively seeing changes brought on by generative AI and CGI advancements.[2] Actors reported being subject to full body scans and other detailed data collection as a condition of employment. At the same time, major movies were using CGI techniques to "revive" dead actors in movies like *The Flash*,

[2]Generative AI in the case of actors is not necessarily a purely generated output, but the use of generative AI assistive tools in replicating actors.

$$\text{Raw Materials} \xrightarrow{\text{Industrial Labor + Technology}} \text{Tangible Commodity}$$

Figure 2.2: Depiction of the traditional understanding of production. Figure from [27]

$$\underset{\text{(Creative labor / CL}^1\text{)}}{\text{Ideas}} \xrightarrow{\text{Technology + Creative Labor / CL}^2} \text{Intellectual Property}$$

Figure 2.3: Depiction of the creative production model from Pang [27]. In this paper, CL1 is called ideative labor, and CL2 is called generative labor. Figure from [27]

giving a vision of what studios could do with actor's data in the future. [26]

2.2 Theoretical Background

2.2.1 A Model of Creative Labor

In its simplest terms, production can be characterized as taking raw materials and applying labor and capital (machinery, technology, etc.), which results in an exchangeable commodity (depicted in 2.2). However, creative labor does not precisely hold to this paradigm. Creative labor takes "raw material" in the form of ideas and results in intellectual property, such as screenplays, movies, set designs, sound mixing, performances, etc., with labor and capital (laptops, cameras, sound equipment, a stage, etc.) still having a role. [27]

Ideas are a non-rival "raw material" in that the formation of ideas does not exhaust any limited resource. However, while human creativity is, in this sense, limitless, it does not mean there is no labor or skill involved in ideation. Training, collaborating, experiencing, discussing, and plenty of other actions go into both building skills in ideation and performing it. In fact, even the act of creating is a key part of ideation. In this way, there is a dialectical relationship between ideation and the production of intellectual property. [27]

To synthesize these ideas, Pang [27] proposes a model of creative labor that takes into account both the labor needed to produce ideas, which I will refer to as ideative labor, and the transformative labor that produces intellectual property, referred to here as generative labor, as depicted in Figure 2.3.

2.2.2 Operaismo and Modern Labor

Labor and tech are often talked about in frameworks of extraction, where technology extracts data from a largely uninvolved workforce. In a broader sense, the relationship between capital and labor is described similarly. However, in "Machinic dispossession and augmented despotism: Digital work in an Amazon warehouse" [1], the author proposes a framework that examines technology's dependence on labor. The paper is based on the operaismo school of thought, which comes from the Autonomist Marxism movement in 1960s Italy. The central idea of operaismo is that "capital depends on labor and needs to secure [labor's] collaboration in production processes in the face of workers' unruliness." [1] From this premise, operaismo theorizes that technology in capitalist organizations serves two separate but closely connected purposes: to enforce the internal hierarchy of firms and to minimize capital's need for labor.

Delfanti, studying labor in an Amazon fulfillment center, observes three critical features of the relationship between labor, digital technologies, and managerial practices and structures. At the core of this framework is machinic dispossession, which is the transformation of labor into fixed capital structures. As opposed to an extractive model of data, machinic dispossession highlights the deprivation of knowledge from workers via technological means. Workers' decision-making and autonomy are "datafied and incorporated in machinery," such warehouse layouts optimized to the point where only machines can comprehend them. [1]

Machinic dispossession enables two interconnected practices, as outlined by Delfanti: algorithmic management and augmented despotism. Algorithmic management consists of labor-management forms enabled by digital technologies, characterized by increasing workload and precarity for workers (i.e., faster pace, higher turnover) to the firm's advantage. The labor process is subject to "standardization and taskification"[1], enabled by machinic dispossession's capture of knowledge such that workers are more controllable and replaceable.

Augmented despotism enables machinic dispossession and algorithmic management by coercing workers' participation in these processes. Augmented despotism consists of the technologically motivated and enabled surveillance and discipline of workers, which can come in the form of gamification, certain conditions of employment, opaque targets, and data collection that only managers can observe among many technologies and practices. [1]

2.3 Application of Theory

As outlined earlier, detailed data collected from streaming services has enabled algorithmic management and machinic dispossession in film and television production. The imperative for more content and the appearance of derisking has led to studio control over detailed decisions in production that would previously be left to workers. For writers, this machinic dispossession is reflected in the mini-room/short-term writer's room structures that have proliferated in Hollywood, as detailed earlier. These new structures remove opportunities for workers to deepen their skills in generative labor, particularly since they no longer stay on throughout the production of a show, which also affects the development of ideative skills since they are closely related. Instead, this skill is now purportedly captured and replaced with statistical approximations based on what data suggests will be successful. Using Pang [27]'s model of creative labor, this data-driven decision-making diminishes the need for ideative labor from creatives, instead using unprecedentedly detailed data to attempt to capture and systematize those skills, in other words, machinic dispossession.

Delfanti's concept of machinic dispossession can be extended to the leverage workers lost during the transition to streaming, as discussed above. What was once leverage to bargain better collective bargaining agreements and cut better individual contracts became solely observable by streaming services. Instead of being used to come to a common understanding about the value and direction of production, data was now warehoused and used for internal data-driven decision-making at an unprecedented level of detail.

In this way, algorithmic management's exemplary characteristics of increased, closely managed workloads and more precarious employment have come through in Hollywood production. Executives feel they can rely more on data-driven decision making in production decisions. As such, writers have less of a role to play, and can easily be managed through more precarious labor structures that enable studios to gain more control over their labor and cut costs. Workers have compared these new conditions to the gig economy created by tech firms' aggressive algorithmic management techniques. [3]

The fight between studios and workers over generative AI can be understood directly as resistance to machinic dispossession, as generative AI takes the outputs of workers' creative labor and transforms them into machinery that performs a statistical approximation of the creative process. Generative AI purports to change the creative process into a more traditional productive process with raw materials coming from scripts and actors' voices and bodies. In this way, generative AI facilitates the transformation of writers' and actors' labor into fixed capital. Unlike the past accumulations of intellectual property, which were treated as financial assets, generative AI turns intellectual property into productive capital (e.g. machinery or technology) that can purportedly be used without the labor that created it. In Pang [27]'s labor model, generative AI targets generative labor, and instead of complimenting it like most technology, it purportedly replaces it.

However, for writers, as noted above, generative AI cannot convincingly replicate the output of professional writers. Instead, generative AI's use is more about legitimizing the technology. Writers were concerned about the studio's disciplinary power in forcing generative AI use. Take, for example, credits on written scripts. Having a generative AI model receive formal credit for writing a script would have two functions: lowering the residuals going to the writer(s) and legitimating the generative AI as a replacement for human labor, even if the writers put just as much work into improving the model's paltry outputs.

For actors, augmented despotism manifests more directly through making the collection of detailed physical data a condition of employment. For background actors in particular, generative AI poses a real threat of replacement, even if the results are noticeably worse and lack the nuances that background actors can bring with their skill. [28]

Actors are facing the machinic dispossession of not only their skill and labor power but also their image and physical qualities. While there are some actors popular enough to negotiate better deals for themselves, most actors and background actors do not have the leverage to individually negotiate against these conditions of employment and prevent generative AI from subsuming the entire creative labor process.

2.4 Contract Resolutions

SAG's and WGA's collective bargaining agreements attempted to create policies that improved their members' working conditions and autonomy. Both guilds largely followed similar tactics to handle structural changes caused by streaming; however, they diverged on generative AI-related issues.

Both unions made significant gains in their new contract with the AMPTP concerning the changing structures of Hollywood. The base compensation rates used in residuals were significantly increased. Also, foreign streaming service residuals were changed to scale with the amount of subscribers. However, the most notable change to streaming-content compensation was the addition of viewership-based streaming bonuses for both TV shows and movies, which is a percentage of the base residual amount. In both the SAG and WGA agreements, the bonus is triggered when 20% or more of a streaming service's subscriber base views a piece of content within 90 days of release or within the first 90 days of each following year. These wins represent an adaptation to a streaming-centric market and will ideally create stability for increasingly sporadic work in Hollywood. [29], [30]

Streamers also agreed to disclose the total number of hours streamed for original programming, a major change in practice that creates a hard metric for evaluating a show's success. This contract provision was an especially difficult fight for the unions but incredibly important for building leverage for future bargaining. As Oscar-winning filmmaker and TV creator Steven Soderbergh put it during the strike: "It's impossible to figure out what's fair if you don't know what's real.... They will not open the books, so how do you figure out what's fair if you don't know what's going on¿' [15] Forcing the studios to "open the books" on streaming gives workers data to point to when bargaining over future contract provisions, as opposed to having to take studios at their word. [29], [30]

Along with income, the WGA agreement also significantly changed staffing practices for TV writer's rooms. The mini-room was eliminated and replaced with minimum staffing levels for shows, depending on their length. Movie writers also won a similar benefit through a guaranteed (in most cases) "second step," meaning that writers had to be kept on for at least one script rewrite. [29] With these wins, writers are reconstructing the pathways for advancement and increasing the opportunities to build skills in creative labor.

While SAG and WGA ended up with similar deals regarding streaming's structural changes, the unions diverged significantly in handling generative AI. WGA's contract focuses on giving writers autonomy in interacting with generative AI. At its core, the WGA CBA's policy on generative AI is that writers cannot be forced to use generative AI or its outputs. Writers can choose to use generative AI with the studio's consent and following any relevant studio policies. Also, studios must disclose if any materials given to the writer are AI-generated. These policies allow writers to legitimize the technology on their terms instead of studios coercing writers into making generative AI seem commonplace. [29]

Studios are also restricted from using generative AI to generate and/or rewrite material for scripts, and any generative AI models that are used in the writing process cannot receive writing credits. On a practical level, these rules mean that generative AI cannot be used to undermine writers' pay, which is determined by credits. These rules around generative AI credit work in concert with the previously mentioned provisions to preserve worker autonomy, and they also specifically target the potential augmented despotism that generative AI could facilitate when used as a disciplinary tool by managers, further preventing studios from forcing through the acceptance of the technology.

One area where the contract is not as prescriptive is regarding generative AI training materials. The contract explicitly leaves the use of writers' scripts as training material unresolved, reserving the right for the WGA or its members to argue that the contract or the law prohibits the "exploitation" (as put by the WGA) of writers' materials. [29] While scripts as training materials still become capital to be accumulated, they can't replace labor due to the other provisions in the contract, thereby disrupting the transformation of labor into fixed capital, as discussed above.

The SAG deal is more complex and detailed in part because it is more permissive in how companies can use generative AI. The SAG contract creates several digital replication/generative AI use tiers and outlines definitions, consent, guidelines, and compensation for each type. Broadly, the contract deals with digital replicas, which depict a specific performer, and synthetic performers, which are created from an amalgamation of performers. Broadly speaking, the contract sets some base expectations, such as actors having to consent each time a digital replica is used and getting paid for equivalent time (with a notable exception, noted below), but otherwise relies on negotiations between individual actors and studios as to the terms of use.[30]

On digital replicas, while consent is required, SAG leadership has explicitly noted that studios can make consent for data collection and/or data use as a condition of employment for actors. In most cases, individual actors do not have the leverage to negotiate, so studios can simply find actors who will consent to data collection and/or data use. These provisions could lead to studios using existing employment structures to create blanket digital replica deals. For example, actors can be paid under Schedule F, which means they get flat pay of at least $80k for a production, no matter how long, instead of a daily or weekly rate. The contract extends this structure to digital replicas of Schedule F actors, meaning replica use would not be compensated on equivalent time. Commentators note that this practice could lead to contracts with blanket replica rights for a franchise or series.[31]

Given that the SAG CBA allows actors to be subject to data collection and digital replication as a condition of employment, actors are less protected against methods of algorithmic despotism. Unless actors can individually negotiate out of it, studios will be able to secure actors' collaboration in building and legitimizing generative AI technologies as a stand-in for creative labor. Ultimately, these provisions allow studios to turn actors' labor into productive capital through data collection.

As for synthetic performers, the contract permits generative AI outputs that are used for performance (i.e., non-background) roles, but they have to be negotiated with SAG. If, the features of specific actors are referenced in the studio's inputs for a generative AI system (e.g. Actor A's eyes and Actor B's nose), consent must also be gathered from those actors. [30] However, beyond negotiations with SAG, the provision would be hard to enforce. Generative AI in its current state tends to "plagiarize" by outputting close if not exact copies of training materials. [32] Still, the contract's enforcement mechanism would only come into effect if an actor's name was specifically used in prompting the generative AI system for the synthetic actor. Even if this high burden of proof is met, actors can only receive monetary restitution. As such, actors have no mechanism to control how their data is used. [33]

It's more difficult to analyze the potential outcomes from the synthetic performer provisions of the contract because they depend on how SAG acts in bargaining. On one extreme, SAG could, in theory, block all synthetic performers, although most likely at the cost of significant bargaining leverage on other issues and there's been no rhetoric from SAG indicating this course of action. Conversely, SAG could allow all synthetic performers, potentially creating additional oversight mechanisms that could have an unknowable effect.

Ultimately, regarding AI, the SAG contract creates a notice and consent regime that largely requires actors to bargain as individuals against large studios. This structure contrasts with the WGA contract, which creates policies to protect the labor and autonomy of writers, which is better suited to protect workers and the creative labor process by the terms outlined in this paper.

Chapter 3

Media on Media: Reasoning for Experiment and Hypothesis

News media has a vital role to play in a strike because of its influence on public opinion. Public support for a strike can boost worker morale and even lend material support through donations to a strike fund or by refusing to do business with struck employers (i.e., not "crossing a picket line")[34] Studies have shown that negative news coverage of a strike can negatively impact perceptions of labor actions, unions, and union members more generally. [35], [36]

During the SAG-AFTRA and WGA strikes, trade publications such as Variety, Deadline, and the Hollywood Reporter were criticized for casting the unions in a negative light [37]. In particular, headlines like "How a WGA Hit Squad Is Shutting Down Hollywood One Shoot at a Time" [38] and social media posts such as "This year's Tonys won't be televised on June 11 after the Writers Guild of America denied a request for a strike waiver from the show's producers." [39]. Some noted that the parent company of all three publications, Penske Media, is a joint owner of an AMPTP-member studio Dick Clark Productions. [40] This sentiment about trade publications reached the point of being parodied by critics of the AMPTP. [41]

At the same time, newsroom workers face pressure from generative AI. Media companies are laying off record numbers of workers [42], [43], newsrooms are already incorporating generative AI to produce news stories [44], and workers' job performance is subject to how social media companies deliver their work. [45] Yet, journalism on technology has had to overcome an overwhelmingly naive approach toward technology firms. [46]

Given the complexity of the relationships between Hollywood labor, studios, technology, and journalism, was there any effect on how news media covered the technology-related demands of the WGA and SAG-AFTRA strikes?

3.1 Methodology

The goal of this quantitative analysis is to determine if there is a quantifiable difference in how the 2023 WGA and SAG-AFTRA strikes were reported on when examined by publisher, by source type (e.g., newspaper, magazine, digital), by topic, and over time, with the goal of uncovering relationships between media institutions themselves and how the strike's relationship to technology was reported on. The hypothesis is that there is a quantifiable difference based on at least one of these groupings.

3.1.1 Data Collection

News articles were sourced from ProQuest from a variety of databases. Using the ProQuest search engine, all articles that specifically mentioned "SAG-AFTRA" or "WGA" were collected. From there, the TDM Studio environment was used to narrow the corpus further. To focus specifically on the 2023 strikes, only articles published in 2023 were included.

Articles were included only if they talked substantively about technology and the strikes. On technology, articles had to mention general terms (e.g., "AI" or "artificial intelligence") at least twice or more technical terms (e.g., "generative AI") at least once. Similarly, to determine if an article was also substantively talking about the strike, an article had to mention one of the two unions once and mention the word "strike" at least twice to be included in the corpus.

After removing any duplicate articles and any documents containing many unrelated news articles, the corpus ended up with 1810 articles from 247 publications across print magazines, web resources, and newspapers. A publication and its online wing are treated separately ("The New York Times" and "The New York Times (Online)"). Still, any republished articles have been removed as part of deduplication, with the earliest published article kept in the corpus.

3.1.2 Quantification

Sentiment Analysis

Articles were quantified using the TextBlob library in Python to generate polarity and subjectivity values for each article. [47] TextBlob uses lexicon-based sentiment analysis to generate the values, meaning that it has a predetermined value for a word (based on its context) and calculates the average value over a document. Polarity measures how positive or negative the wording of an article is and then generates a value in a range of [-1.0, 1.0], with 1 being positive, -1 being negative, and 0 being neutral. Subjectivity works similarly by assigning a value to each document based on the subjective language used on a range of [0,1.0], with 0 representing factual language (e.g. "X happens because") and 1 representing opinionated language (e.g. "I think that X happens because").

Important to note is that, being lexicon-based, the accuracy of these measures will be limited in relation to the meaning of "polarity" or "subjectivity." In other words, there might not be much meaning to a document's polarity or subjectivity score or even to an average score of document groups within the corpus. However, the scores can still show if there are any significant differences in wording between documents to find differences between groups. For example, a group of documents with a much higher subjectivity score may not be more "subjective," but it shows a quantifiable difference to be further explored. More advanced sentiment analysis techniques were considered, but the small size of the corpus limited options.

Sentence Embedding Clusters

Article text was tokenized into sentences using the English-language Punkt sentence tokenizer in the Natural Language Toolkit Python Library [48] producing $\tilde{7}$3k sentences. Sentences were also labeled as technology-related and/or union-related based on keywords in the sentences. All of the sentences were then embedded using the all-MiniLM-L6-v2 model [49]. The all-MiniLM-L6-v2 model was selected because it produces 384-dimensional (low-dimensional relative to the dataset), dense vectors, making it well-suited for clustering tasks.

The DBSCAN clustering algorithm was used to find clusters in the embedding space to determine if any natural clusters coincided with the document groupings. [50] DBSCAN works by finding clusters of dense, separated areas of data points and labeling outliers outside of those clusters as noise. DBSCAN was chosen because of its ability to find clusters of arbitrary shapes (unlike K-means, which assumes clusters are spherical) and to filter out outliers in the embedding space. DBSCAN also does not need to be provided an assumed number of clusters.

The DBSCAN's parameters, ϵ and *minPts*, were initially set using standard heuristics. [51] *MinPts*, the minimum amount of points needed for a cluster, was initially set to $2 * \dim$ and ϵ, the maximum distance between two points to be considered as neighbors, was set by finding the "knee" in an ordered set of the average distance to the nearest *minPts* number of points using the "Kneedle" algorithm to find the optimal point [52].

3.1.3 Topic Extraction

Term Frequency-Inverse Document Frequency (TF-IDF) was used to group articles by topic. TF-IDF determines the topics of articles by calculating how frequently a term occurs in a document (Term Frequency) and penalizing it by how often a term appears in the corpus (Inverse Document Frequency). This calculation finds words that are specifically important to a document, which approximates the topic of a document. The TF-IDF values can then be used to vectorize the documents so they can be clustered by topic. [53]

One drawback to vectorization using TF-IDF is that the generated matrix is incredibly sparse, which can result in poor clustering performance. This limitation is remedied by using Latent Semantic Analysis (LSA), a topic modeling technique, to reduce the matrix's sparsity and dimensionality, thereby making clustering more efficient and effective.

Different vectorization methods using machine learning models such as Doc2Vec and BERT were considered. However, the data set is too small to train or fine-tune a model and is too specialized to reliably use fully pre-trained models due to words falling out of vocabulary.

The documents were grouped using KMeans clustering. The elbow method, a plot of the within-cluster sum of squares, was used to find the k-value. The 'silhouette core, a cluster cohesiveness and separation measure, was used to determine cluster goodness post hoc. All vectorization and clustering methods were performed using the SKLearn Library in Python [54].

3.2 Results

3.2.1 Sentence Embedding Clustering

Following the procedure above, $minPts = 768$ and $\epsilon = 1.1109564$, which resulted in 1 cluster of about 95% points and the rest labeled as outliers. Further parameter changes resulted in points shifting to either the large cluster or noise, but no values tested produced more than a single cluster. This result suggests that, in embedding space, the sentences do not densely group together anywhere within the space. The outlined process was also performed for only the technology-related sentences, union-related sentences, and the union of the two sets, all of which had the same result.

3.2.2 Sentiment Analysis

The polarity and subjectivity were graphed as a daily average and a rolling 30-day average. As shown by the charts below, there is no trend within the data related to time (Figure A.1). Similar to the temporal data, when the sentiment analysis values are averaged by publication (Figure A.7) and source type (Figure A.3), there is no significant difference from the overall mean.

As described above, TF-IDF was used to vectorize the documents, and LSA was used to reduce the dimensionality of the matrix to 500 components, preserving 68.5% of the variance of the original TF-IDF matrix. The number of components, a hyperparameter for LSA, was determined through experimentation to find the value where the rate of increase of variance explained slowed down.

Through the elbow plot (Figure A.4), $k = 4$ (number of clusters) was selected as the hyperparameter for K Means clustering. A TSNE projection (Figure A.5) shows the clusters are not well separated. A silhouette score, used to evaluate the goodness of clustering, of 0.019 (on a range of [-1.0, 1.0] also indicates this conclusion.

Examining the word clouds of the clusters (Figure A.6) reveals that the topics covered by each cluster: cluster 1 contains articles about the WGA strike; cluster 2 includes articles about the SAG-AFTRA strike; cluster 3, which is significantly smaller than the other clusters, primarily represents articles from one news source; and cluster 4 focuses on both the strikes as they relate to artificial intelligence.

For each cluster, the average polarity and subjectivity are within one standard deviation of the mean value for the data set. While the clusters cover distinct topics, there is no significant quantifiable difference in the sentiments by topic.

Chapter 4

Conclusion

While the long-term effects of the SAG and WGA strikes and collective bargaining agreements are unknowable at this time, this analysis evaluates the agreements based on the political economy, techno-labor relations present before the strike, and workers' experience-based concerns surrounding changing technology in creative production. A quantitative analysis of news media is also presented to find evidence of media firms' positioning, but the results were inconclusive.

WGA and SAG's actions certainly have put tech-labor relations and the role of unions into focus as more workers face changing conditions due to technology. SAG and WGA set examples of workers and their representatives setting technology policy for themselves to shape their own working conditions. With a languid, apathetic legislature combined with an administrative state under attack from a largely corporate-friendly judiciary, it is all the more important to better understand the collective actions of and perception-setting around labor.

Appendix A

Appendix

Source Type	Amount of Articles in Corpus
Newspapers	1416
Web Resources	347
Magazines	47

Table A.1: Number of Articles in Corpus by Type

	Polarity	Subjectivity
measurement range	[-1.0,1.0]	[0.0,1.0]
mean	.08	.45
standard deviation	.06	.06
min	-.15	.21
25th percentile	.05	.41
50th percentile	.08	.45
75th percentile	.12	.49
max	.33	.66

Table A.2: Descriptive Statistics for Sentiment Measures

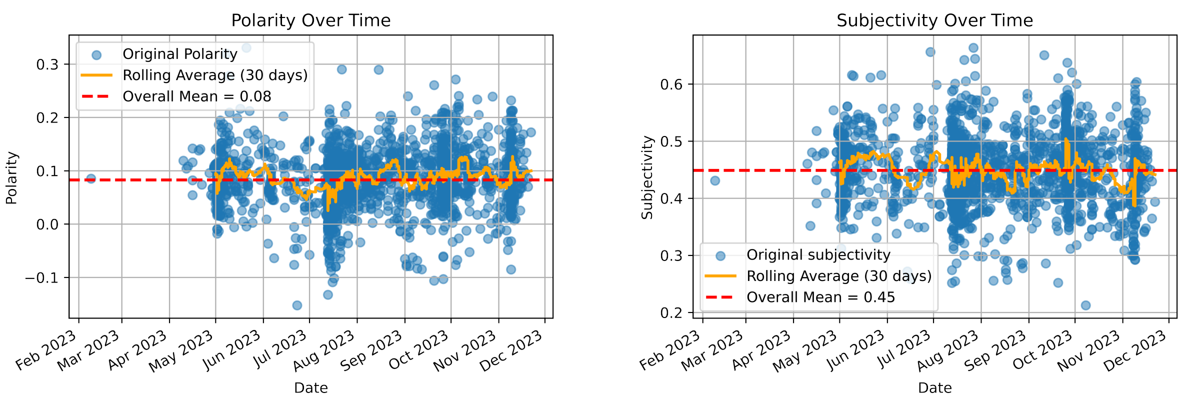


Figure A.1: Sentiment Analysis Values Over Time

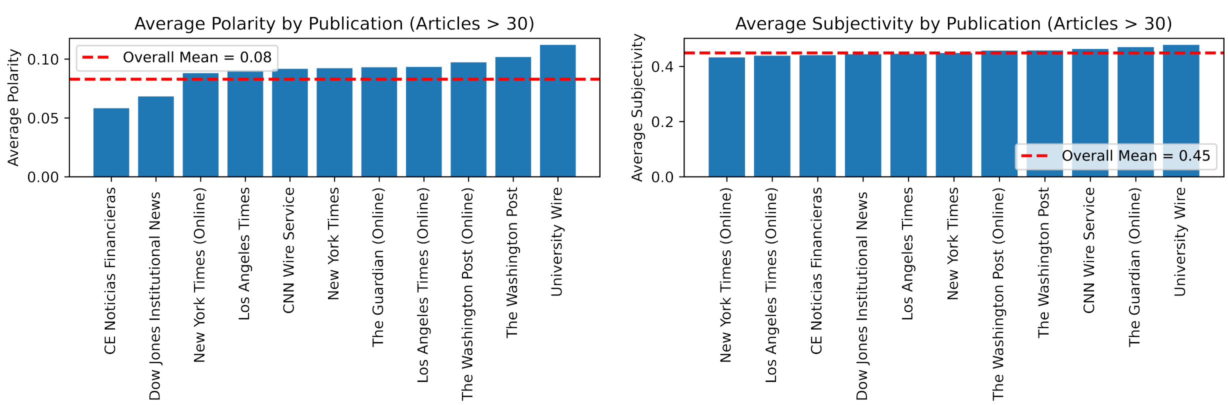


Figure A.2: Average Sentiment Analysis Values By Publication

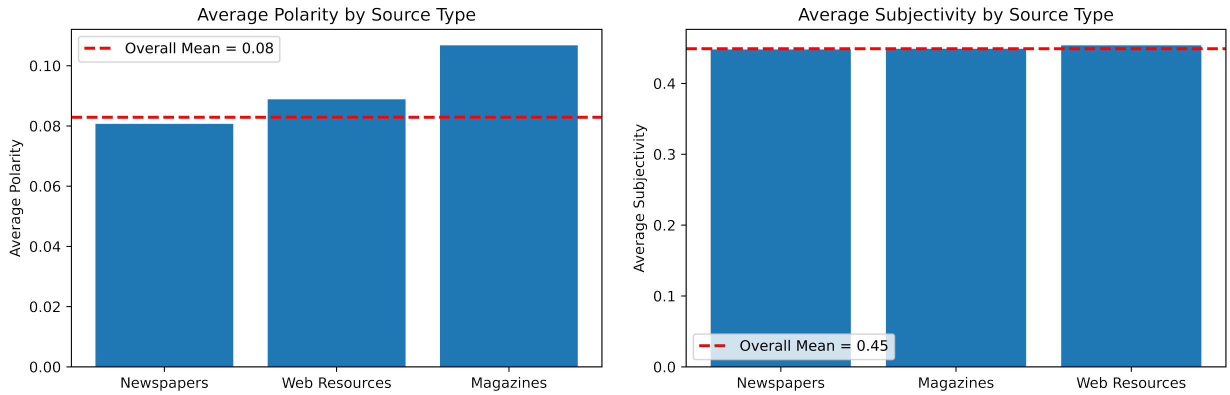


Figure A.3: Average Sentiment Analysis Values By Source

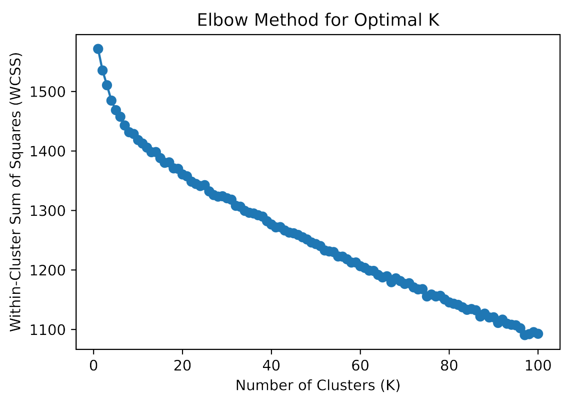


Figure A.4: Elbow Plot

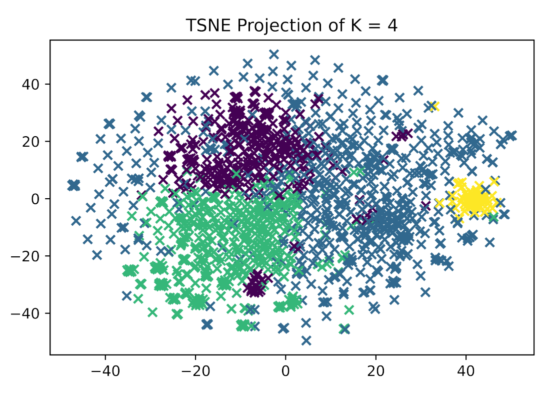
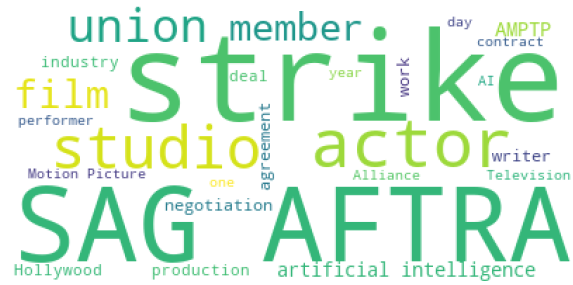


Figure A.5: TSNE Projection

(a) Word Cloud for Cluster 1

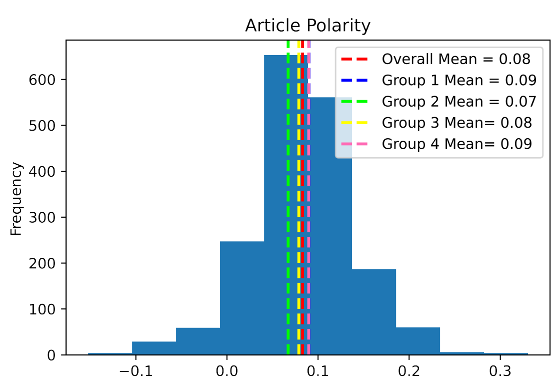


(b) Word Cloud for Cluster 2

(c) Word Cloud for Cluster 3

(d) Word Cloud for Cluster 4

Figure A.6: Word Clouds for each cluster



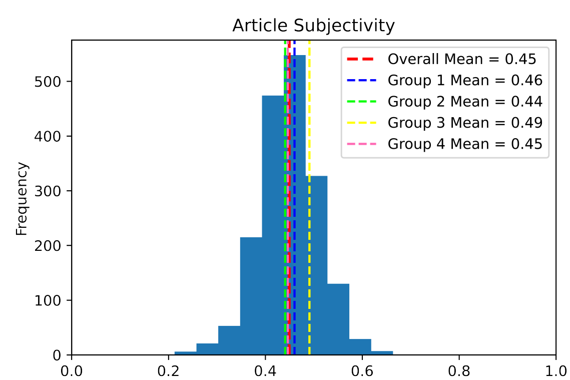


Figure A.7: Histogram with KMeans means

References

[1] A. Delfanti, "Machinic dispossession and augmented despotism: Digital work in an Amazon warehouse," en, *New Media & Society*, vol. 23, no. 1, pp. 39–55, Jan. 2021, Publisher: SAGE Publications, ISSN: 1461-4448. DOI: 10.1177/1461444819891613. [Online]. Available: https://doi.org/10.1177/1461444819891613 (visited on 11/13/2023).

[2] J. Wilson, "Special Effects of Unions in Hollywood," eng, *Loyola of Los Angeles Entertainment Law Journal*, vol. 12, no. 2, pp. 403–440, 1992. [Online]. Available: https://heinonline.org/HOL/P?h=hein.journals/laent12&i=421 (visited on 12/09/2023).

[3] A. Wolf, *The WGA Strike Is a Fight Against Silicon Valley's Gigification of the Entire Economy*, en-US, May 2023. [Online]. Available: https://jacobin.com/2023/05/wga-writers-guild-of-america-strike-silicon-valley-gigification-economy-residuals (visited on 11/21/2023).

[4] L. Weiss, T. Noah, T. Noah, *et al.*, "Did the Actors' Union Really Win?" *The New Republic*, Dec. 2023, ISSN: 0028-6583. [Online]. Available: https://newrepublic.com/article/177422/sag-aftra-union-strike-ai (visited on 03/28/2024).

[5] P. White and D. Patten, *It's Official: WGA Members Overwhelmingly Ratify New Three-Year Deal With Studios*, en-US, Oct. 2023. [Online]. Available: https://deadline.com/2023/10/writers-strike-deal-approved-ratification-vote-1235567930/ (visited on 03/27/2024).

[6] Z. Phillips, *SAG-AFTRA and WGA Strikes: All the Major Dates to Know*, en_US, Dec. 2023. [Online]. Available: https://www.etonline.com/sag-and-wga-strikes-all-the-major-dates-to-know-207915 (visited on 03/27/2024).

[7] A. Liu, "The Domination of Hollywood: Ownership Changes and Present-Day Implications," en, May 2020, Publisher: [object Object]. DOI: 10.26153/TSW/11464. [Online]. Available: https://repositories.lib.utexas.edu/handle/2152/84485 (visited on 03/07/2024).

[8] H. Vanhala, "A critical political economy analysis of the streaming industry's structure," in *Television's Streaming Wars*, Num Pages: 17, Routledge, 2023, ISBN: 978-1-00-333564-1.

[9] PricewaterhouseCoopers, *Perspectives and insights: Global Entertainment and Media Outlook 2023–2027*, en_gx. [Online]. Available: https://www.pwc.com/gx/en/industries/tmt/media/outlook/insights-and-perspectives.html (visited on 03/07/2024).

[10] R. Porter, *Peak TV Climbs Again in 2022, Nearly Reaches 600 Scripted Series*, en-US, Jan. 2023. [Online]. Available: https://www.hollywoodreporter.com/tv/tv-news/peak-tv-2022-count-599-scripted-shows-1235298139/ (visited on 03/07/2024).

[11] J. Kuehn and R. Lampe, "Competition and product composition: Evidence from hollywood," en, *International Journal of Industrial Organization*, vol. 91, p. 102 981, Dec. 2023, ISSN: 01677187. DOI: 10.1016/j.ijindorg.2023.102981. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167718723000620 (visited on 03/07/2024).

[12] A. Barnhardt, *Marvel Studios Executive Addresses MCU Strategy Pivot*, en, Mar. 2024. [Online]. Available: https://comicbook.com/movies/news/marvel-studios-brad-winderbaum-film-tv-strategy-pivot-explained/ (visited on 03/18/2024).

[13] K. Toussaint, *TikTok, K-dramas, and canceled subscriptions: How a writers' strike could look different in the streaming era*, en, May 2023. [Online]. Available: https://www.fastcompany.com/90889321/tiktok-k-dramas-and-canceled-subscriptions-how-a-writers-strike-could-look-different-in-the-streaming-era (visited on 12/09/2023).

[14] C. Littleton, *Writers' Fear of Joining the 'Gig Economy' Fuels WGA Picket Lines*, en-US, May 2023. [Online]. Available: https://variety.com/2023/tv/news/gig-economy-writers-strike-1235608830/ (visited on 11/22/2023).

[15] K. Richardson, *SAG Actors Want Higher Residuals. Here's How Streaming Residuals Work*, en-US, Aug. 2023. [Online]. Available: https://www.rollingstone.com/tv-movies/tv-movie-features/how-streaming-residuals-work-sag-actors-strike-hollywood-1234802623/ (visited on 12/09/2023).

[16] *Demystifying WGA Television Residuals*, en, Oct. 2018. [Online]. Available: https://cdas.com/demystifying-wga-television-residuals/ (visited on 05/10/2024).

[17] *Residuals Survival Guide*, en, Mar. 2022. [Online]. Available: https://www.wga.org/members/finances/residuals/residuals-survival-guide (visited on 05/10/2024).

[18] *Summary of the 2020 WGA MBA*, en. [Online]. Available: https://www.wga.org/contracts/contracts/mba/summary-of-the-2020-wga-mba (visited on 05/10/2024).

[19] *High-Budget SVOD Residuals*. [Online]. Available: https://www.wgaeast.org/levies-and-payments/residuals/high-budget-svod/ (visited on 05/10/2024).

[20] SAG-AFTRA, *Understanding Residuals for SAG-AFTRA Performers*, Dec. 2020. [Online]. Available: https://www.youtube.com/watch?v=p4U7CRtmdVM (visited on 05/10/2024).

[21] C. Mims, "How Netflix's Algorithms and Tech Feed Its Success," en-US, *Wall Street Journal*, Jul. 2023, ISSN: 0099-9660. [Online]. Available: https://www.wsj.com/articles/how-netflixs-algorithms-and-tech-feed-its-success-90632b92 (visited on 12/09/2023).

[22] G. R. Martin, *Writers On Set / Not a Blog*, en-US, May 2023. [Online]. Available: https://georgerrmartin.com/notablog/2023/05/08/writers-on-set/ (visited on 12/09/2023).

[23] G. Maddaus, *How the WGA Decided to Harness — but Not Ban — Artificial Intelligence*, en-US, May 2023. [Online]. Available: https://variety.com/2023/biz/news/wga-ai-writers-strike-technology-ban-1235610076/ (visited on 12/09/2023).

[24] C. Pulliam-Moore, *SAG-AFTRA's new contract hinges on studios acting responsibly with AI*, en, Nov. 2023. [Online]. Available: https://www.theverge.com/2023/11/18/23962349/sag-aftra-tentative-agreement-generative-artificial-intelligence-vote (visited on 12/09/2023).

[25] D. Calacci, "Building Dreams Beyond Labor: Worker Autonomy in the Age of AI," en, *Interactions*, vol. 30, no. 6, pp. 48–51, Nov. 2023, ISSN: 1072-5520, 1558-3449. DOI: 10.1145/3626245. [Online]. Available: https://dl.acm.org/doi/10.1145/3626245 (visited on 03/15/2024).

[26] B. Allyn, "Movie extras worry they'll be replaced by AI. Hollywood is already doing body scans," en, *NPR*, Aug. 2023. [Online]. Available: https://www.npr.org/2023/08/02/1190605685/movie-extras-worry-theyll-be-replaced-by-ai-hollywood-is-already-doing-body-scan (visited on 12/09/2023).

[27] L. Pang, "The Labor Factor in the Creative Economy," en, *Social Text*, vol. 27, no. 2, pp. 55–76, 2009, ISSN: 0164-2472, 1527-1951. DOI: 10.1215/01642472-2008-022. [Online]. Available: https://read.dukeupress.edu/social-text/article/27/2%20(99)/55/33591/The-Labor-Factor-in-the-Creative-EconomyA-Marxist (visited on 03/14/2024).

[28] V. Mayer, "The Production of Extras in a Precarious Creative Economy," en, in *Precarious Creativity*, University of California Press, 2016, ISBN: 978-0-520-96480-8. DOI: 10.1353/book.63390. [Online]. Available: https://muse.jhu.edu/book/63390 (visited on 03/14/2024).

[29] *Summary of the 2023 WGA MBA*, en. [Online]. Available: https://www.wgacontract2023.org/the-campaign/summary-of-the-2023-wga-mba (visited on 11/21/2023).

[30] *Summary of Tentative Agreement*. [Online]. Available: https://www.sagaftra.org/files/sa_documents/TV-Theatrical_23_Summary_Agreement_Final.pdf (visited on 03/25/2024).

[31] W. Cho, *Actors' AI Protections Are a Step Forward, But There's Reason to Worry*, en-US, Nov. 2023. [Online]. Available: https://www.hollywoodreporter.com/business/business-news/where-sag-aftra-deal-may-miss-ai-protections-1235646988/ (visited on 03/30/2024).

[32] G. Marcus and R. Southen, *Generative AI Has a Visual Plagiarism Problem*, en, Jan. 2024. [Online]. Available: https://spectrum.ieee.org/midjourney-copyright (visited on 03/31/2024).

[33] K. Kilkenny, *"A Crazy, Heated Time": SAG-AFTRA Members Debate Deal Merits Ahead of Ratification Vote*, en-US, Dec. 2023. [Online]. Available: https://www.hollywoodreporter.com/business/business-news/sag-aftra-ratification-vote-debate-ai-1235711336/ (visited on 03/30/2024).

[34] E. K. Kelloway, L. Francis, V. M. Catano, and K. E. Dupré, "Third-party support for strike action," *Journal of Applied Psychology*, vol. 93, no. 4, pp. 806–817, 2008, Place: US Publisher: American Psychological Association, ISSN: 1939-1854. DOI: 10.1037/0021-9010.93.4.806.

[35] D. E. Schmidt, "Public opinion and media coverage of labor unions," en, *Journal of Labor Research*, vol. 14, no. 2, pp. 151–164, Jun. 1993, ISSN: 1936-4768. DOI: 10.1007/BF02685662. [Online]. Available: https://doi.org/10.1007/BF02685662 (visited on 12/03/2023).

[36] J. V. Kane and B. J. Newman, "Organized Labor as the New Undeserving Rich?: Mass Media, Class-Based Anti-Union Rhetoric and Public Support for Unions in the United States," en, *British Journal of Political Science*, vol. 49, no. 3, pp. 997–1026, Jul. 2019, Publisher: Cambridge University Press, ISSN: 0007-1234, 1469-2112. DOI: 10.1017/S000712341700014X. [Online]. Available: https://www.cambridge.org/core/journals/british-journal-of-political-science/article/abs/organized-labor-as-the-new-undeserving-rich-mass-media-classbased-antiunion-rhetoric-and-public-support-for-unions-in-the-united-states/6CF19F2860C07ABAF0B7C787D8731A0C (visited on 12/03/2023).

[37] H. Brooks, *How Media Bias Twists Public Perception of the Writers' Strike*, en-US, May 2023. [Online]. Available: https://www.theedgemedia.org/wga-strike-2023/ (visited on 12/01/2023).

[38] P. White, *How A WGA Hit Squad Is Shutting Down Hollywood One Shoot At A Time*, en-US, May 2023. [Online]. Available: https://deadline.com/2023/05/writers-strike-how-a-wga-hit-squad-is-taking-down-hollywood-one-shoot-at-a-time-1235360087/ (visited on 12/02/2023).

[39] *Variety on X: "UPDATE: This year's Tonys won't be televised on June 11 after the Writers Guild of America denied a request for a strike waiver from the show's producers. https://t.co/JNE0iOLqoD" / X*, en, May 2023. [Online]. Available: https://twitter.com/Variety/status/1657179906533638144 (visited on 12/01/2023).

[40] C. Nast, *Writers Strike Coverage in 2007 Was the Nikki Finke Show. Twitter Has Changed the Game*, en-US, Section: tags, May 2023. [Online]. Available: https://www.vanityfair.com/news/2023/05/writers-strike-coverage-2023-nikki-finke-show-twitter (visited on 12/10/2023).

[41] Carol Lombardini [@ItsMeCarolAMPTP], *For some reason THR omitted this part from our interview. how many times do I have to tell these people they're supposed to print EXACTLY what the AMPTP tells them??? https://t.co/2K3rPnmwaa*, en, Tweet, Sep. 2023. [Online]. Available: https://twitter.com/ItsMeCarolAMPTP/status/1706728319075729707 (visited on 12/01/2023).

[42] T. Spangler, *Layoffs Hit Condé Nast and Vox Media*, en-US, Nov. 2023. [Online]. Available: https://variety.com/2023/digital/news/layoffs-conde-nast-vox-media-1235815307/ (visited on 12/10/2023).

[43] E. Washburn and T. Roush, *2023 Media Layoffs: CBC Cuts 10% Of Workforce*, en, Section: Business, Dec. 2023. [Online]. Available: https://www.forbes.com/sites/emilywashburn/2023/12/05/2023-media-layoffs-cbc-cuts-10-of-workforce/ (visited on 12/10/2023).

[44] M. Sato, *CNET is overhauling its AI policy and updating past stories*, en, Jun. 2023. [Online]. Available: https://www.theverge.com/2023/6/6/23750761/cnet-ai-generated-stories-policy-update (visited on 12/10/2023).

[45] C. Warzel, *The Great Social Media–News Collapse*, en, Section: Technology, Nov. 2023. [Online]. Available: https://www.theatlantic.com/technology/archive/2023/11/social-media-news-readership-decline/675890/ (visited on 12/10/2023).

[46] B. Wallace, *Inside the Nasty Battle Between Tech and Journalism*, en, May 2021. [Online]. Available: https://nymag.com/intelligencer/2021/05/tech-vs-journalism-silicon-valley-reporters.html (visited on 12/10/2023).

[47] S. Loria, "Textblob documentation," *Release 0.15*, vol. 2, 2018.

[48] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit.* " O'Reilly Media, Inc.", 2009.

[49] *Sentence-transformers/all-MiniLM-L6-v2 · Hugging Face*, Mar. 2024. [Online]. Available: https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2 (visited on 03/11/2024).

[50] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96, Portland, Oregon: AAAI Press, Aug. 1996, pp. 226–231. (visited on 03/11/2024).

[51] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN," en, *ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, Sep. 2017, ISSN: 0362-5915, 1557-4644. DOI: 10.1145/3068335. [Online]. Available: https://dl.acm.org/doi/10.1145/3068335 (visited on 03/11/2024).

[52] V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan, "Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior," en, in *2011 31st International Conference on Distributed Computing Systems Workshops*, Minneapolis, MN, USA: IEEE, Jun. 2011, pp. 166–171, ISBN: 978-1-4577-0384-3. DOI: 10.1109/ICDCSW.2011.20. [Online]. Available: http://ieeexplore.ieee.org/document/5961514/ (visited on 03/11/2024).

[53] P. Bafna, D. Pramod, and A. Vaidya, "Document clustering: TF-IDF approach," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Mar. 2016, pp. 61–66. DOI: 10.1109/ICEEOT.2016.7754750. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7754750 (visited on 12/08/2023).

[54] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.